

EXHIBIT 6

to the Attorney Declaration



Library Genesis: Miner's Hut
You need to login to gain access to all forum sections.

Search...



Quick links FAQ

Notifications Private messages **sallysue** ▼

[Board index](#) < [Library Genesis Forum](#) < [Law & politics: copyright, piracy, open access, legal aspects](#)

Copyright law of the United States

Post Reply



Search this topic...



4 posts • Page **1** of **1**

Copyright law of the United States



by **Librarian** » 14 May 2023, 13:58

Librarian

Library Genesis staff

Posts: **2260**

Joined: 12 Jun 2010, 13:22

Here is a selection of some essential reading to provide a better understanding of the threat posed by the U.S. authorities to pirate/shadow libraries and their enthusiasts.

Wikipedia: [Copyright law of the United States](#)

Copyright Law of the United States (Title 17) and Related Laws Contained in Title 17 of the United States Code

<https://www.copyright.gov/title17/>

Criminal Copyright Infringement -- 17 U.S.C. 506(a) And 18 U.S.C. 2319

<https://www.justice.gov/archives/jm/criminal-resource-manual-1847-criminal-copyright-infringement-17-usc-506a-and-18-usc-2319>

Copyright Infringement -- Second Element -- Infringement

<https://www.justice.gov/archives/jm/criminal-resource-manual-1849-copyright-infringement-second-element-infringement>

Copyright Infringement -- Third Element -- Willfulness

<https://www.justice.gov/archives/jm/criminal-resource-manual-1850-copyright-infringement-third-element-willfulness>

Copyright Infringement -- Fourth Element -- Commercial Advantage Or Private Financial Gain

<https://www.justice.gov/archives/jm/criminal-resource-manual-1851-copyright-infringement-fourth-element-commercial-advantage-or>

Copyright Infringement -- Penalties -- 17 U.S.C. 506(a) And 18 U.S.C. 2319

<https://www.justice.gov/archives/jm/criminal-resource-manual-1852-copyright-infringement-penalties-17-usc-506a-and-18-usc-2319>

"Prosecuting Intellectual Property Crimes" Manual

<https://www.justice.gov/file/442151/download> (PDF) (via [CCIPS Documents And Reports](#))

Legal considerations when gathering online cyber threat intelligence and purchasing data from illicit sources

<https://www.justice.gov/criminal-ccips/page/file/1252341/download> (PDF)

Re: Copyright law of the United States

by **IHaveDoubts** » 27 May 2023, 19:06



IHaveDoubts

Posts: 4
Joined: 27 May 2023,
18:49

Hello, I have a doubt: is it really illegal (or should it be) to download books from here or other shadow library? I knew Library Genesis existed a time ago, but I registered here today. Don't know if I would do the right thing downloading stuff this way, I don't want to be a criminal, I don't want to steal money from others, I just need knowledge so I could grow as a person. But there are a lot of barriers to do so and others are not helping that much abusing copyright laws instead of doing things for the vulnerable population getting books easier. How I can be better at life if knowledge is for a tiny group of people? So I guess I need to talk this with someone.

Re: Copyright law of the United States

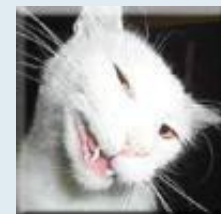
by **dehydrator** » 28 May 2023, 02:21



IHaveDoubts

LG is not a pirate resource. Piracy means profit. We contribute to a more even, free and fair distribution and preservation of knowledge and educational information throughout the world.

All this is reflected in the LG code. The topic is intended, first of all, to observe elementary network security, because the "sharks" of the so-called copyright are very aggressive



dehydrator

Posts: 5348
Joined: 15 Jan 2010,
07:39

Re: Copyright law of the United States

by **IHaveDoubts** » 29 May 2023, 16:08



IHaveDoubts

Posts: 4
Joined: 27 May 2023,
18:49

dehydrator wrote: ↑

28 May 2023, 02:21

IHaveDoubts

LG is not a pirate resource. Piracy means profit. We contribute to a more even, free and fair distribution and preservation of knowledge and educational information throughout the world.

All this is reflected in the LG code. The topic is intended, first of all, to observe elementary network security, because the "sharks" of the so-called copyright are very aggressive

Then I would support LG. It would be great to attack misinformations surrounding this: there are people, as I said before, that would attack LG and similar stuff "in the name of law" when abuse of this law/copyright and greed exist.



Quick Reply

Subject:

Re: Copyright law of the United States

Full Editor & Preview

Submit

Post Reply



4 posts • Page **1** of **1**

[Return to "Law & politics: copyright, piracy, open access, legal aspects"](#)

[Board index](#)

[Contact us](#)

[The team](#)

[Members](#)

[Delete cookies](#)

All times are UTC-04:00

Powered by [phpBB®](#) Forum Software © phpBB Limited

[Privacy](#) | [Terms](#)

phpBB® Library Genesis: Miner's Hut
creating communities You need to login to gain access to all forum sections.

Search...

Quick links FAQ Notifications Private messages salty sue

Board index Library Genesis Forum Content hoarding, book scanning, file processing

[help needed][content scraping] Digitalia Hispánica 11 posts • Page 1 of 1

Post Reply Search this topic...

[help needed][content scraping] Digitalia Hispánica
by **kruppload** » 20 Nov 2023, 16:32
<https://www.digitaliapublishing.com/w/9...hispanica->
Hello all, Digitalia Hispánica claims to have "more than 50,000 full text titles from around 400 publishers from Spain and Latin America" at the link above. This is not the largest database, but it has a large number of academic & literary works in Spanish that aren't already available in the other shadow libraries, to my knowledge.

Getting an ebook requires 2 levels of account access: the website has to be viewed through a subscribing institution, which can access a subset of the full Digitalia Hispánica catalog. Then, a "personal" account (throwaway email accepted) is needed to download an ebook, which uses the .acsm/Adobe Digital Editions system & can be de-DRMed similarly to Internet Archive.

As with my last thread, I am not sure about the technical requirements or if "scraping" is the right word to describe systematic downloading from this database, but wanted to bring it to the attention of those who might.

kruppload
Posts: 168
Joined: 16 Feb 2020, 20:34

Re: [help needed][content scraping] Digitalia Hispánica
by **Librarian** » 21 Nov 2023, 11:12
We are looking for someone with an access to *Digitalia Hispánica* from within your academic institution's network!
As we understand, the subscribers/licensees are mostly various Hispanic universities.

Librarian
Library Genesis staff
Posts: 2323
Joined: 12 Jun 2010, 13:22

Re: [help needed][content scraping] Digitalia Hispánica
by **rabbitabi** » 22 Nov 2023, 17:54
I am interested in working on this. Let's talk.

rabbitabi
Posts: 1
Joined: 22 Nov 2023, 17:52

Re: [help needed][content scraping] Digitalia Hispánica
by **nozomi** » 22 Nov 2023, 18:18

Librarian wrote: ↑ 21 Nov 2023, 11:12
We are looking for someone with an access to *Digitalia Hispánica* from within your academic institution's network!
As we understand, the subscribers/licensees are mostly various Hispanic universities.

I have already contacted OP few days ago since one of my friends is working on spanish resources. If OP can join either discord or TG it will be much easier for us.

nozomi
Posts: 978
Joined: 08 Oct 2015, 06:56

Re: [help needed][content scraping] Digitalia Hispánica
by **Librarian** » 22 Nov 2023, 23:55

rabbitabi wrote: ↑ 22 Nov 2023, 17:54
I am interested in working on this. Let's talk.

Thanks for joining us!
Do you have an academic access to Digitalia Hispánica? Could you download a few sample .acsm files from there?

nozomi wrote: ↑ 22 Nov 2023, 18:18
I have already contacted OP few days ago since one of my friends is working on spanish resources. If OP can join either discord or TG it will be much easier for us.

As a compromise solution for everyone, we can just discuss the task here, since neither Discord nor Telegram are completely anonymous and secure, while the forum also provides greater transparency of the process, so that more people can also learn something.

Librarian
Library Genesis staff
Posts: 2323
Joined: 12 Jun 2010, 13:22

Re: [help needed][content scraping] Digitalia Hispánica
by **Cegred** » 23 Nov 2023, 09:45
(NB: this is rabbitabi. I just had my old account in my computer and was using my phone when I wrote the previous rabbitabi message because I could not access this account)

Thank you for the welcome, librarian. Happy to engage in conversation here.

As noted by nozomi above, I have been working on / trying to work on some Spanish collections: to stay on track I'll limit myself to Digitalia Hispanica here.

From my end, things seem to be different as from OP's. My institutional access requires registration with institutional email, from the same institution as the access itself. Moreover, acsm downloads are capped to 10 per week, and returning the items does not remove this limit.

I could provide proxies to access Digitalia Hispanica: however the issue remains that institutional emails are required to download the acsms. The most promising path for a large scale scrap is, in my humble opinion, to develop a script that rebuilds the file via the online visor. However, this visor does not rely on jpg files but rather a 'background jpg' plus some text that appears on top. You can see a bit of the html here <https://imgur.com/a/fDCu4x6>

However, this is punching above my weight and we really would need the support of a developer: not only for this platform but for many others (mostly Elibro & UNEbook)

Cegred
Posts: 12
Joined: 13 Nov 2022, 09:29

Re: [help needed][content scraping] Digitalia Hispánica
by **Librarian** » 24 Nov 2023, 01:54

Cegred wrote: ↑ 23 Nov 2023, 09:45
My institutional access requires registration with institutional email, from the same institution as the access itself. Moreover, acsm downloads are capped to 10 per week, and returning the items does not remove this limit.

Librarian
Library Genesis staff
Posts: 2323
Joined: 12 Jun 2010, 13:22

That's not good news. 😞

As you may understand, most hope lies with this method of scraping the content to get native ebooks with the least effort. Do you know anything about whether higher limits are ever possible? Perhaps you know someone at other institutions who could check with their subscriptions?

Could you provide a sample page (save it as an HTML file with your desktop browser) from which any book (.acsm file) is actually downloaded? Perhaps we can find some vulnerabilities in their system.

🗨️ Cegred wrote: 1

23 Nov 2023, 09:45

The most promising path for a large scale scrap is, in my humble opinion, to develop a script that rebuilds the file via the online visor. However, this visor does not rely on jpg files but rather a 'background.jpg' plus some text that appears on top. You can see a bit of the html here <https://imgur.com/a/fDCu4x6>

If they don't display a full image of each page, it makes little sense to deal with scraping that way, since reconstructing a PDF from such data would be extremely cumbersome and produce ugly results.

🗨️ Cegred wrote: 1

23 Nov 2023, 09:45

However, this is punching above my weight and we really would need the support of a developer: not only for this platform but for many others (mostly Elibro & UNEbook)

If you need any help with tools development or researching something, you may always ask here by creating separate topics and describing the task, or, if the matter is too sensitive, just [PM me](#). There are several developers around who can help with scraping, automation and other stuff.

Re: [help needed][content scraping] Digitalia Hispánica

by Cegred » 25 Nov 2023, 14:16

Thank you for your thorough response.This is the html from the site at the point of download (i.e. just before clicking on [https://www-digitaliapublishing-com.\[redacted\].idm.oclc.org/downloadadobedrmInk/pdf](https://www-digitaliapublishing-com.[redacted].idm.oclc.org/downloadadobedrmInk/pdf) to trigger the download of the acsm file)

<https://pastebin.com/nCP4q1Rh>

While I agree that reconstructing the pdf is a subpar option, finding a way to crack this visors would open the doors to a number of platforms that rely on those. You may find an example for the html code of the visor displaying a random page of a book here <https://pastes.io/eot26uwag7>

update: please check the html for the visor, as it seems that the pictures may lie there in base64.

Re: [help needed][content scraping] Digitalia Hispánica

by OMEGAR0JO » 25 Dec 2023, 01:04

Tengo Acceso a la pagina

Re: [help needed][content scraping] Digitalia Hispánica

by Cegred » 25 Dec 2023, 17:22

Anyone who could look into this visor / base64 thing? We haave access!

Re: [help needed][content scraping] Digitalia Hispánica

by woolbeat » 20 Jan 2024, 18:02

🗨️ Cegred wrote: 1

25 Nov 2023, 14:16

Thank you for your thorough response.This is the html from the site at the point of download (i.e. just before clicking on [https://www-digitaliapublishing-com.\[redacted\].idm.oclc.org/downloadadobedrmInk/pdf](https://www-digitaliapublishing-com.[redacted].idm.oclc.org/downloadadobedrmInk/pdf) to trigger the download of the acsm file)

<https://pastebin.com/nCP4q1Rh>

While I agree that reconstructing the pdf is a subpar option, finding a way to crack this visors would open the doors to a number of platforms that rely on those. You may find an example for the html code of the visor displaying a random page of a book here <https://pastes.io/eot26uwag7>

update: please check the html for the visor, as it seems that the pictures may lie there in base64.

FIY, the

CODE: SELECT ALL

data:image/png;base64

images are just icons. One of them is "EN", another shows a red "Z" and the other 2 are symbols. You can check it out yourself here: <https://base64.guru/converter/decode/image> . Just paste the contents of the src attribute, e.g.

CODE: SELECT ALL

data:image/png;base64,iVBORwCODE_REMOVED_FOR_BREVITY_ggg==

A network request log (.har) would be useful. Here's a guide how to capture it: <https://support.google.com/admanager/an...85977hl=en> . Chrome, Firefox, Edge, Brave, they all have this functionality.

Just remember to open the devtools before you load any document, preferably from the link to the document. Likely the javascript is loading everything dynamically, so you won't see it in the HTML source.

The .har file itself is just a JSON with a different extension. You can rename it, edit out the information you consider sensitive, then upload it to pastebin like you did before. No request contents are saved, so it will be a small file.

Quick Reply

Subject: Re: [help needed][content scraping] Digitalia Hispánica

Full Editor & Preview

Submit

Post Reply

11 posts • Page 1 of 1

< Return to "Content hoarding, book scanning, file processing"

Board index

Contact us

The team

Members

Delete cookies

All times are UTC-04:00

https://forum.mhut.org/viewtopic.php?f=10&t=18692&p=202617#p202617 2/3

